

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 00:51:44 ; Search time 69 Seconds
(without alignments)
145.736 Million cell updates/sec

Title: US-09-980-263-1

Perfect score: 147

Sequence: 1 GMOGPAGSGWEGSGSPGVTFPLFSP 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	26	4 AAG62388	AAG62388 Alternati
2	147	100.0	26	4 AAB48914	AAB48914 Human ace
3	147	100.0	26	4 AAB50032	AAB50032 Acetylcho
4	147	100.0	26	5 AAU98024	AAU98024 Human rea
5	147	100.0	31	2 AAW74588	AAW74588 Amino aci
6	147	100.0	31	2 AAW68146	AAW68146 Human ACH
7	147	100.0	37	4 AAB50035	AAB50035 Acetylcho
8	147	100.0	53	4 AAB50036	AAB50036 Acetylcho
9	147	100.0	53	5 AAB50037	AAB50037 GFP-fused
10	147	100.0	600	2 AAW48797	AAW48797 Human ace
11	63	42.9	145	7 ADC33366	ADC33366 Human nov
12	61.5	41.8	575	4 ABB11475	ABB11475 Human R31
13	61	41.5	98	3 AAB51865	AAB51865 Human sec
14	60	40.8	54	2 AAR37744	AAR37744 Collagen
15	60	40.8	54	2 AAR93255	AAR93255 Collagen
16	60	40.8	334	8 ADL91154	ADL91154 Mouse fic
17	60	40.8	334	8 ADL91152	ADL91152 Human fic
18	60	40.8	633	2 AAR37745	AAR37745 Collagen
19	60	40.8	633	2 AAR93257	AAR93257 Collagen
20	60	40.8	633	2 AAR57655	AAR57655 Collagen
21	60	40.8	1065	2 AAR37745	AAR37745 Collagen
22	60	40.8	1065	2 AAR93256	AAR93256 Collagen
23	60	40.8	1065	2 AAR57654	AAR57654 Collagen
24	60	40.8	1466	4 AAE02534	AAE02534 Bovine al
25	60	40.8	1466	4 AAE02533	AAE02533 Bovine al

26	59	40.1	532	2 AAW40114	AAW40114 Human alp
27	58	39.5	226	7 AD564162	AD564162 Human pro
28	58	39.5	251	6 ADA54929	ADA54929 Human pro
29	58	39.5	274	8 ADQ66581	ADQ66581 Novel hum
30	58	39.5	393	5 AAU86144	AAU86144 Human PRO
31	58	39.5	393	6 ADA54950	ADA54950 Human PRO
32	58	39.5	393	7 ADC31141	ADC31141 Human nov
33	58	39.5	393	7 ADJ37315	ADJ37315 Human tum
34	58	39.5	393	8 ADG68239	ADG68239 Human PRO
35	58	39.5	393	8 ADQ65780	ADQ65780 Novel hum
36	58	39.5	421	7 ADF74163	ADF74163 Human nov
37	58	39.5	635	4 AAM78798	AAM78798 Human pro
38	58	39.5	638	5 ABJ01025	ABJ01025 Human bre
39	58	39.5	638	6 ABU69146	ABU69146 Human NOV
40	58	39.5	638	8 ADO08295	ADO08295 Human NOV
41	58	39.5	644	4 AAM79782	AAM79782 Human pro
42	58	39.5	703	6 ABP96315	ABP96315 Human col
43	58	39.5	717	6 ABP96314	ABP96314 Human col
44	58	39.5	733	6 ABU69145	ABU69145 Human NOV
45	58	39.5	733	8 ADO08293	ADO08293 Human NOV

ALIGNMENTS

RESULT 1
AAG62388
ID AAG62388 standard; peptide; 26 AA.
XX
AC AAG62388;
XX
DT 31-AUG-2001 (first entry)
XX
DE Alternatively splice AChE product C-terminus AChE-R.
XX
KW Antisense oligonucleotide; acetylcholine esterase; AChE; dystonia;
KW cholinergic neurotransmission; progressive neuromuscular disorder;
KW myasthenia gravis; Eaton-Lambert disease; muscular dystrophy; PTSD;
KW amyotrophic lateral sclerosis; post-traumatic stress disorder;
KW multiple sclerosis; post-stroke sclerosis; post-injury muscle damage;
KW excessive re-innervation.
XX
OS Unidentified.
XX
FN WO200136627-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-IL000763.
XX
PR 16-NOV-1999; 99IL-00132972.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Soreq H, Seidman S;
XX
PT WPI; 2001-336003/35.
XX
DR New antisense oligonucleotide targeted to acetylcholine esterase mRNA,
XX useful for treating or preventing progressive neuromuscular disorders
XX such as myasthenia gravis.
XX
PT Disclosure; Fig 1; 124pp; English.
XX
CC Sequences AAW44810 - AAW44822 represent antisense oligonucleotides
CC targeting the acetylcholine esterase (AChE) mRNA. AChE is involved in the
CC termination of cholinergic neurotransmission, by hydrolysing the
CC neurotransmitter acetylcholine. Mammalian AChE is encoded by one gene but
CC alternative splicing at its 3' end yields three different mRNA
CC transcripts which encode protein with distinct carboxyl termini. All
CC three proteins are catalytically active. AChE has morphogenic, non-
CC catalytic capacities too. AChE antisense oligonucleotides are used in
CC treating or preventing a progressive neuromuscular disorder. Examples of

disorders which are treatable using the antisense oligonucleotides include myasthenia gravis, Eaton-Lambert disease, muscular dystrophy, amyotrophic lateral sclerosis, post-traumatic stress disorder (PTSD), multiple sclerosis, dystonia, post-stroke sclerosis, post-injury muscle damage, excessive re-innervation and post-exposure to AChE inhibitors. The present sequence represents the C-terminus of an alternatively spliced AChE gene product

XX SQ Sequence 26 AA;

Query Match 100.0%; Score 147; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGVTPLFSP 26
DB 1 GMOGPAGSGWEGSGSPPGVTPLFSP 26

RESULT 2
ID AAB48914 standard; peptide; 26 AA.
XX AAB48914;
XX 16-MAR-2001 (first entry)
XX Human acetylcholinesterase (AChE) C-terminal peptide, SEQ ID NO:1.
DE Acetylcholinesterase; AChE; readthrough peptide; ARP; splice variant;
XX human; epitope; C-terminal peptide; antibody; central nervous system;
KW CNS stress; psychological insult; physical insult; chemical insult;
KW blood-brain barrier disruption; elevated glucocorticoid level;
KW Alzheimer's disease; diagnosis.
XX Homo sapiens.
XX WO200073343-A2.
XX 07-DEC-2000.
XX 31-MAY-2000; 2000WO-IL000312.
XX 31-MAY-1999; 99IL-00130225.
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Soreq H, Kauffer D, Friedman A, Seidman S;
XX WPI; 2001-061514/07.
XX Antibody specific to acetylcholinesterase or its C-terminal peptide derivative useful for diagnosing, ventral nervous system stress, elevated glucocorticoid level, disruption of blood-brain barrier and Alzheimer's disease.
XX Claim 4; Page 42; 44pp; English.
XX The invention relates to antibodies which recognise acetylcholinesterase (AChE) or a C-terminal peptide thereof (particularly AAB48914-B48916). The AChE splice variant, AChE-R, and AChE-R mRNA, have been found to be elevated in response to central nervous system (CNS) insults. The invention therefore also relates to a method for diagnosing CNS stress, and also elevated glucocorticoid levels, disruption of the blood-brain barrier or Alzheimer's disease using a sample (e.g., serum or cerebrospinal fluid) and an antibody of the invention. The CNS stress which may be diagnosed using the antibodies is preferably that caused by psychological insult, physical insult (head injury, head trauma, or exposure to irradiation) or chemical insult (exposure to insecticide or nerve gas). The present sequence represents a human AChE C-terminal peptide (termed AChE readthrough peptide (ARP) in the specification), which is specifically claimed as an epitope which is recognised by an antibody of the invention

XX SQ Sequence 26 AA;

Query Match 100.0%; Score 147; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGVTPLFSP 26
DB 1 GMOGPAGSGWEGSGSPPGVTPLFSP 26

RESULT 3
ID AAB50032 standard; peptide; 26 AA.
XX AAB50032;
XX 14-MAR-2001 (first entry)
XX Acetylcholinesterase readthrough peptide ARP-1.
DE ARP-1; haemostatic; acetylcholinesterase; AChE; cell growth; human;
KW cell differentiation; thrombocytopenia; post-irradiation condition;
KW post-chemotherapy condition; blood loss; stress-induced male infertility.
XX Homo sapiens.
XX WO200073427-A2.
XX 07-DEC-2000.
XX 31-MAY-2000; 2000WO-IL000311.
XX 31-MAY-1999; 99IL-00130224.
XX 02-SEP-1999; 99IL-00131707.
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Soreq H, Eldor A, Deutch V, Grisaru D;
XX WPI; 2001-061523/07.
XX New regulatory peptides having cell growth and cell differentiation activity derived from the C-terminal region of acetylcholinesterase useful in promoting growth, survival and differentiation of stem cells.
XX Claim 8; Page 50; 133pp; English.
XX The present sequence is a C-terminal peptide of acetylcholinesterase (AChE). This peptide is acetylcholinesterase "readthrough" peptide (ARP-1). This peptide has a cell growth and/or cell differentiation activity. The peptide may be used in ex vivo or in vivo expansion of haematopoietic stem cells and neural progenitors, and in the promotion of megakaryocytic differentiation of haematopoietic stem cells. In addition, the present peptide may be used in for promoting expansion of committed neural progenitors in a developing embryo, in cultured embryonic stem cells, and embryoid bodies derived from them. The present peptide may further be used in the treatment of thrombocytopenia, post-irradiation conditions, post-chemotherapy conditions, and conditions following massive blood loss, in inducing synthesis of AChE mRNA, and in promoting formation of hematom bodies. Antibodies directed against the present peptide are useful for diagnosing stress-induced male infertility

XX SQ Sequence 26 AA;

Query Match 100.0%; Score 147; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGVTPLFSP 26
DB 1 GMOGPAGSGWEGSGSPPGVTPLFSP 26

```

Db      1 GMOGPAGSGWEEGSGPPGVTPLFSP 26
|||||
RESULT 5
AAW74588
ID AAW74588 standard; protein; 31 AA.
XX
XX AAW74588;
AC
XX
XX 21-DEC-1998 (first entry)
DT
XX
XX Amino acid sequence of the human AChE variant 7.
DE
XX Nuclease resistance; inhibition; human; acetyl-cholinesterase; AChEs;
KW central nervous system; CNS.
KW
XX Homo sapiens.
OS
XX W09839486-A1.
XX PN
XX 11-SEP-1998.
PD
XX
XX 06-MAR-1998; 98WO-US004503.
XX PF
XX 06-MAR-1997; 97US-0040203P.
XX PR
XX (YISS ) YISSUM RES & DEV CO.
PA
XX (KOHN/) KOHN K I.
PA
XX Soreq H, Seidman S, Shohami E;
PI
XX WPI; 1998-506377/43.
XX DR
XX Treatment of injury to central nervous system - by administration of
PT inhibitor of acetylcholinesterase production.
XX
XX Disclosure; Page 62; 88pp; English.
XX
XX This is the amino acid sequence of a human acetyl-cholinesterase (AChE)
CC variant used in the method of the invention, where inhibitors of AChE
CC used to treat injury to the central nervous system (CNS). The AChE
CC inhibitor can also be used to facilitate transplantation of neuronal
CC cells to the CNS of a patient. The inhibitor can also be used to improve
CC hippocampal neuron survival following injury to the CNS. The CNS injury
CC that can be treated with the method include epilepsy, stroke,
CC Huntington's disease, head injury, spinal injury, pain, Parkinson's
CC disease, myelin deficiencies, neuromuscular disorders, neurological
CC amyotrophic lateral sclerosis, Alzheimer's disease, and affective
CC disorders of the brain
XX
XX Sequence 31 AA;
SQ
Query Match 100.0%; Score 147; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps
QY 1 GMOGPAGSGWEEGSGPPGVTPLFSP 26
DB 6 GMOGPAGSGWEEGSGPPGVTPLFSP 31
|||||
RESULT 6
AAW68146
ID AAW68146 standard; protein; 31 AA.
XX
XX AAW68146;
AC
XX
XX 05-OCT-1998 (first entry)
DT
XX Human AChE splice variant E1-4-I4-E5.
DE
XX Nuclease resistant; acetylcholinesterase; human; myasthenia gravis;
KW

```

KW Parkinson's disease; Alzheimer's disease; central nervous system;
KW neuromuscular junction; cholinergic signalling; brain.

OS Homo sapiens.

PN WO9826062-A2.

XX 18-JUN-1998.

XX 12-DEC-1997; 97WO-US023598.

XX 12-DEC-1996; 96US-003526GP.

PR 13-FEB-1997; 97US-003777FP.

PR 02-MAY-1997; 97US-00850347.

PR 21-JUL-1997; 97US-0053334P.

XX (YISS) YISSUM RES & DEV CO.

PA (KOHN/) KOHN K I.

PI Soreq H, Seidman S, Eckstein F, Friedman A, Kaufer D;

XX WPI; 1998-348522/30.

XX Synthetic nuclease resistant antisense oligodeoxynucleotides - directed

PT against acetylcholinesterase, useful for treating Parkinson's and

PT Alzheimer's diseases and myasthenia gravis.

XX Disclosure; Fig 12; 89pp; English.

XX This represents the amino acid sequence of a human acetylcholinesterase

GC (AChE) splice variant. The invention provides sequences shown in AAV41278

CC to AAV41285 that represent synthetic nuclease resistant antisense

CC oligodeoxynucleotides which are capable of selectively modulating human

CC acetylcholinesterase (AChE) production. These oligonucleotides are

CC targeted to a splice junction in a splice variant of AChE mRNA and are

CC capable of selectively modulating human AChE production in the central

CC nervous system and neuromuscular junction. The invention also provides a

CC method for determining the efficacy of these human AChE specific

CC antisense oligonucleotides. These antisense oligonucleotides can be used

CC to restore balanced cholinergic signalling in the brain, particularly

CC related to learning and memory as well as stress disorders, Parkinson's

CC and Alzheimer's disease. They can also be used to reduce production and

CC therefore deposition of AChE in the neuromuscular junctions of patients

CC with e.g. myasthenia gravis. The oligonucleotides work effectively at low

CC doses while avoiding many of the side effects associated with tacrine and

CC related cholinergic drugs for Alzheimer's disease and pyridostigmine and

XX related drugs for myasthenia gravis

XX Sequence 31 AA;

SQ Query Match 100.0%; Score 147; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. No. 1.6e-10;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

Db 6 GMOGPAGSGWEEGSGPPGVTPFLFSP 31

RESULT 7

AAB50035

ID AAB50035 standard; peptide; 37 AA.

XX ARP-2; haemostatic; acetylcholinesterase; AChE; cell growth; human;

XX cell differentiation; thrombocytopenia; post-irradiation condition;

XX post-chemotherapy condition; blood loss; stress-induced male infertility.

XX Homo sapiens.

OS WO200073427-A2.

XX 07-DEC-2000.

XX 31-MAY-2000; 2000WO-IL000311.

XX 31-MAY-1999; 99IL-00130224.

XX 02-SEP-1999; 99IL-00131707.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Soreq H, Eldor A, Deutch V, Grisaru D;

XX WPI; 2001-061523/07.

XX New regulatory peptides having cell growth and cell differentiation

PT activity derived from the C-terminal region of acetylcholinesterase

PT useful in promoting growth, survival and differentiation of stem cells.

XX Example 10; Page 76; 133pp; English.

XX The present invention relates to C-terminal peptides of

CC acetylcholinesterase (AChE) (see AAB50032-B50034). The peptides of the

CC present invention have cell growth and/or cell differentiation activity.

CC The peptides may be used in ex vivo or in vivo expansion of

CC haematopoietic stem cells and neural progenitors, and in the promotion of

CC megakaryocytic differentiation of hematopoietic stem cells. In addition,

CC the peptides may be used in for promoting expansion of committed neural

CC progenitors in a developing embryo. The peptides may further be used in

CC embryoid bodies derived from them. The peptides may further be used in

CC the treatment of thrombocytopenia, post-irradiation conditions, post-

CC chemotherapy conditions, and conditions following massive blood loss, in

CC inducing synthesis of AChE mRNA, and in promoting formation of hematol

CC bodies. Antibodies directed against the peptides are useful for

CC diagnosing stress-induced male infertility. The present sequence is a C-

CC terminal AChE "readthrough" peptide (ARP-2)

XX Sequence 37 AA;

SQ Query Match 100.0%; Score 147; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

Db 12 GMOGPAGSGWEEGSGPPGVTPFLFSP 37

RESULT 8

AAB50036

ID AAB50036 standard; protein; 53 AA.

XX AAB50036;

XX 14-MAR-2001 (first entry)

XX Acetylcholinesterase protein #1 used in a yeast two-hybrid system.

XX ARP; haemostatic; acetylcholinesterase; AChE; cell growth; human;

OS Homo sapiens.

PN WO200073427-A2.

XX 07-DEC-2000.

XX 31-MAY-2000; 2000WO-IL000311.

XX 31-MAY-1999; 99IL-00130224.

XX 02-SEP-1999; 99IL-00131707.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Soreq H, Eldor A, Deutch V, Grisaru D;

XX WPI; 2001-061523/07.

XX New regulatory peptides having cell growth and cell differentiation

PT activity derived from the C-terminal region of acetylcholinesterase

PT useful in promoting growth, survival and differentiation of stem cells.

XX Example 10; Page 76; 133pp; English.

XX The present invention relates to C-terminal peptides of

CC acetylcholinesterase (AChE) (see AAB50032-B50034). The peptides of the

CC present invention have cell growth and/or cell differentiation activity.

CC The peptides may be used in ex vivo or in vivo expansion of

CC haematopoietic stem cells and neural progenitors, and in the promotion of

CC megakaryocytic differentiation of hematopoietic stem cells. In addition,

CC the peptides may be used in for promoting expansion of committed neural

CC progenitors in a developing embryo. The peptides may further be used in

CC embryoid bodies derived from them. The peptides may further be used in

CC the treatment of thrombocytopenia, post-irradiation conditions, post-

CC chemotherapy conditions, and conditions following massive blood loss, in

CC inducing synthesis of AChE mRNA, and in promoting formation of hematol

CC bodies. Antibodies directed against the peptides are useful for

CC diagnosing stress-induced male infertility. The present sequence is a C-

CC terminal AChE "readthrough" peptide (ARP-2)

XX Sequence 37 AA;

SQ Query Match 100.0%; Score 147; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

Db 12 GMOGPAGSGWEEGSGPPGVTPFLFSP 37

RESULT 8

AAB50036

ID AAB50036 standard; protein; 53 AA.

XX AAB50036;

XX 14-MAR-2001 (first entry)

XX Acetylcholinesterase protein #1 used in a yeast two-hybrid system.

XX ARP; haemostatic; acetylcholinesterase; AChE; cell growth; human;

XX cell differentiation; thrombocytopenia; post-irradiation condition;

XX post-chemotherapy condition; blood loss; stress-induced male infertility.

XX Homo sapiens.

OS WO200073427-A2.

PN WO200073427-A2.

XX 07-DEC-2000.

XX 31-MAY-2000; 2000WO-IL000311.

XX 31-MAY-1999; 99IL-00130224.

XX 02-SEP-1999; 99IL-00131707.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Soreq H, Eldor A, Deutch V, Grisaru D;

XX WPI; 2001-061523/07.

XX New regulatory peptides having cell growth and cell differentiation

PT activity derived from the C-terminal region of acetylcholinesterase

PT useful in promoting growth, survival and differentiation of stem cells.

XX Example 10; Page 76; 133pp; English.

XX The present invention relates to C-terminal peptides of

CC acetylcholinesterase (AChE) (see AAB50032-B50034). The peptides of the

CC present invention have cell growth and/or cell differentiation activity.

CC The peptides may be used in ex vivo or in vivo expansion of

CC haematopoietic stem cells and neural progenitors, and in the promotion of

CC megakaryocytic differentiation of hematopoietic stem cells. In addition,

CC the peptides may be used in for promoting expansion of committed neural

CC progenitors in a developing embryo. The peptides may further be used in

CC embryoid bodies derived from them. The peptides may further be used in

CC the treatment of thrombocytopenia, post-irradiation conditions, post-

CC chemotherapy conditions, and conditions following massive blood loss, in

CC inducing synthesis of AChE mRNA, and in promoting formation of hematol

CC bodies. Antibodies directed against the peptides are useful for

CC diagnosing stress-induced male infertility. The present sequence is a C-

CC terminal AChE "readthrough" peptide (ARP-2)

XX Sequence 37 AA;

SQ Query Match 100.0%; Score 147; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

Db 12 GMOGPAGSGWEEGSGPPGVTPFLFSP 37

RESULT 8

AAB50036

ID AAB50036 standard; protein; 53 AA.

XX AAB50036;

XX 14-MAR-2001 (first entry)

XX Acetylcholinesterase protein #1 used in a yeast two-hybrid system.

XX ARP; haemostatic; acetylcholinesterase; AChE; cell growth; human;

XX cell differentiation; thrombocytopenia; post-irradiation condition;

XX post-chemotherapy condition; blood loss; stress-induced male infertility.

XX Homo sapiens.

OS WO200073427-A2.

PN WO200073427-A2.

XX 07-DEC-2000.

XX 31-MAY-2000; 2000WO-IL000311.

XX 31-MAY-1999; 99IL-00130224.

XX 02-SEP-1999; 99IL-00131707.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Soreq H, Eldor A, Deutch V, Grisaru D;

XX WPI; 2001-061523/07.

XX New regulatory peptides having cell growth and cell differentiation

PT activity derived from the C-terminal region of acetylcholinesterase

PT useful in promoting growth, survival and differentiation of stem cells.

XX Example 10; Page 76; 133pp; English.

XX The present invention relates to C-terminal peptides of

CC acetylcholinesterase (AChE) (see AAB50032-B50034). The peptides of the

CC present invention have cell growth and/or cell differentiation activity.

CC The peptides may be used in ex vivo or in vivo expansion of

CC haematopoietic stem cells and neural progenitors, and in the promotion of

CC megakaryocytic differentiation of hematopoietic stem cells. In addition,

CC the peptides may be used in for promoting expansion of committed neural

CC progenitors in a developing embryo. The peptides may further be used in

CC embryoid bodies derived from them. The peptides may further be used in

CC the treatment of thrombocytopenia, post-irradiation conditions, post-

CC chemotherapy conditions, and conditions following massive blood loss, in

CC inducing synthesis of AChE mRNA, and in promoting formation of hematol

CC bodies. Antibodies directed against the peptides are useful for

CC diagnosing stress-induced male infertility. The present sequence is a C-

CC terminal AChE "readthrough" peptide (ARP-2)

XX Sequence 37 AA;

SQ Query Match 100.0%; Score 147; DB 4; Length 37;

Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

Db 12 GMOGPAGSGWEEGSGPPGVTPFLFSP 37

RESULT 8

AAB50036

ID AAB50036 standard; protein; 53 AA.

XX AAB50036;

XX 14-MAR-2001 (first entry)

XX Acetylcholinesterase protein #1 used in a yeast two-hybrid system.

XX ARP; haemostatic; acetylcholinesterase; AChE; cell growth; human;

XX cell differentiation; thrombocytopenia; post-irradiation condition;

XX post-chemotherapy condition; blood loss; stress-induced male infertility.

XX Homo sapiens.

OS WO200073427-A2.

PR 02-SEP-1999; 99IL-00131707.
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX PI Soreq H, Eldor A, Deutch V, Grisaru D;
 XX WPI; 2001-061523/07.
 DR
 PT New regulatory peptides having cell growth and cell differentiation
 PT activity derived from the C-terminal region of acetylcholinesterase
 PT useful in promoting growth, survival and differentiation of stem cells.
 XX
 XX PS Claim 8; Page 87; 133pp; English.
 XX
 CC The present invention relates to C-terminal peptides of
 CC acetylcholinesterase (ACHE) (see AAB50032-B50034). The peptides of the
 CC present invention have cell growth and/or cell differentiation activity.
 CC The peptides may be used in ex vivo or in vivo expansion of
 CC haematopoietic stem cells and neural progenitors, and in the promotion of
 CC megakaryocytic differentiation of hematopoietic stem cells. In addition,
 CC the peptides may be used in for promoting expansion of committed neural
 CC progenitors in a developing embryo, in cultured embryonic stem cells, and
 CC embryoid bodies derived from them. The peptides may further be used in
 CC the treatment of thrombocytopenia, post-irradiation conditions, post-
 CC chemotherapy conditions, and conditions following massive blood loss, in
 CC inducing synthesis of ACHE mRNA, and in promoting formation of hematoin
 CC bodies. Antibodies directed against the peptides are useful for
 CC diagnosing stress-induced male infertility. The present sequence is a C-
 CC terminal ACHE "readthrough" protein (ARP), which was used in a yeast two-
 CC hybrid system, to screen for ARP binding partners
 XX
 SQ Sequence 53 AA;
 Query Match 100.0%; Score 147; DB 4; Length 53;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GMOQPGAGSGWEEGSGSPGVTPLFSP 26
 |||||
 Db 28 GMOQPGAGSGWEEGSGSPGVTPLFSP 53
 |||||
 RESULT 9
 ABG31331
 ID ABG31331 standard; protein; 53 AA.
 AC ABG31331;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE GFP-fused ACHE variant expression construct, pGARP related protein.
 XX
 KW Nervous system; drug assay; acetylcholinesterase; ACHE; brain;
 KW isoform variance; ACHE blocker; muscarinic receptor; M1; M2;
 KW pyridostigmine; muscarinic receptor blocker; scopolamine;
 KW M1 receptor blocker; pirenzepine; anxiety; post-traumatic stress;
 KW Alzheimer's disease; muscle malfunctioning; neurodegenerative disorder;
 KW xenobiotic damage; panic; neuromuscular disorder; Parkinson's disease;
 KW Huntington's chorea; muscle fatigue; multiple chemical sensitivity;
 KW autism; multiple sclerosis; Sjogren's disease; GFP; pGARP;
 KW green fluorescent protein.
 XX
 OS Unidentified.
 XX
 XX WO200240994-A2.
 PN
 XX 23-MAY-2002.
 ED
 XX 14-NOV-2001; 2001WO-IL001051.
 XX
 XX 14-NOV-2000; 2000US-0247970P.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA

XX Soreq H, Meshorer E, Sklan E, Shoham S;
 XX WPI; 2002-490152/52.
 DR
 XX
 PT Evaluating effect of drugs on nervous system by comparing effect of drug
 PT on acetylcholinesterase, ACHE activity in brain of test animal following
 PT challenge by ACHE blocker and comparing it with control group.
 XX
 XX PS Example; Page 52; 114pp; English.
 XX
 CC The present invention relates to a method and system for evaluating an
 CC effect on the nervous system of a test drug. The method comprises
 CC comparing the effect of the drug on acetylcholinesterase (ACHE) catalytic
 CC activity or isoform variance in a brain of a test animal following a
 CC challenge by an ACHE blocker or a blocker of ACHE and muscarinic
 CC receptors M1 and M2 (e.g. pyridostigmine) and comparing this effect with
 CC that of a known agent, preferably a non-selective muscarinic receptor
 CC blocker (e.g. scopolamine) or a specific M1 receptor blocker (e.g.
 CC pirenzepine). The method is useful for evaluating an effect on the
 CC nervous system of a test drug, including drugs for the treatment of
 CC anxiety conditions, post-traumatic stress, Alzheimer's disease, muscle
 CC malfunctioning, neurodegenerative disorders, damage resulting from
 CC exposure to xenobiotics, panic, neuromuscular disorders, Parkinson's
 CC disease, Huntington's chorea, muscle fatigue, multiple chemical
 CC sensitivity, autism, multiple sclerosis and Sjogren's disease. The
 CC present sequence represents a protein described in relation to green
 CC fluorescent protein (GFP)-fused ACHE variant expression construct pGARP
 CC in the examples of the present invention
 XX
 SQ Sequence 53 AA;
 Query Match 100.0%; Score 147; DB 5; Length 53;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GMOQPGAGSGWEEGSGSPGVTPLFSP 26
 |||||
 Db 28 GMOQPGAGSGWEEGSGSPGVTPLFSP 53
 |||||
 RESULT 10
 AAW48797
 ID AAW48797 standard; protein; 600 AA.
 AC AAW48797;
 XX
 DT 07-OCT-1998 (first entry)
 XX
 DE Human acetylcholine esterase-I4 readthrough splice variant.
 XX
 KW Human acetylcholine esterase-I4 readthrough splice variant; CNS;
 KW blood/brain barrier; BBB; I4 peptide; antibiotic; brain tumour; glioma;
 KW chemotherapeutic drug; central nervous system.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Region 1..574
 FT /note= "This region is encoded by exons 1-4 of ACHE"
 FT Region 575..599
 FT /note= "I4 peptide encoded by intron 4 of ACHE; this
 FT sequence is claimed by the inventors under claim 2 in the
 FT specification"
 FT Region 600
 FT /note= "residue encoded by exon 5 of ACHE"
 FT
 XX WO9822132-A2.
 PN
 XX 28-MAY-1998.
 PD
 XX 20-NOV-1997; 97WO-US021696.
 PF
 XX

PR 20-NOV-1996; 96US-0031194P.
 PR 12-DEC-1996; 96US-0035266P.
 PR 21-JUL-1997; 97US-0053200P.
 XX
 PA (YISS) VISSUM RES & DEV CO.
 PA (KOHN/) KOHN K I.
 XX
 PI Soreq H, Friedman A, Seidman S, Kaufer D;
 XX WPI; 1998-312172/27.
 XX
 PT Increasing the permeability of the blood/brain barrier - using e.g.
 PT adrenaline, atropine or acetylcholine esterase I4 splice variant peptide,
 PT useful for imaging and/or treatment of central nervous system disorders.
 XX
 XX Claim 1, 2; Page 45; 71pp; English.
 XX
 CC The present sequence represents human acetylcholine esterase-I4 (ACH-E-I4)
 CC readthrough splice variant. The protein sequence comprises residues
 CC encoded by exons 1-4 of human ACH-E followed by residues encoded by intron
 CC 4, while the last residue of the protein is encoded by exon 5 of ACH-E.
 CC The invention provides a pharmaceutical composition, for facilitating
 CC passage of compounds through the blood/brain barrier (BBB), comprising of
 CC ACH-E-I4, I4 peptide or ACH-E-I4 analogues together with a pharmaceutically
 CC acceptable carrier. The pharmaceutical composition is claimed to
 CC facilitate a reversible disruption of the BBB allowing transport of
 CC compounds through the BBB. The compounds, e.g. imaging agents,
 CC antibiotics or chemotherapeutic drugs, are claimed to be useful for the
 CC diagnosis and treatment of diseases or disorders of the CNS such as
 CC infections, neurochemical disorders, brain tumours, gliomas, etc
 XX
 SQ Sequence 600 AA;
 Query Match 100.0%; Score 147; DB 2; Length 600;
 Best Local Similarity 100.0%; Pred. No. 3.2e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GMQPGAGSGWEEGSGSPGVTPLFSP 26
 Db 575 GMQPGAGSGWEEGSGSPGVTPLFSP 600
 RESULT 11
 ADC33366
 ID ADC33366 standard; protein; 145 AA.
 XX
 AC ADC33366;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3448.
 XX
 KW Human; diagnosis; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 XX 24-SEP-2002; 2002WO-US030474.
 PF
 XX 24-SEP-2001; 2001US-0324631P.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX WPI; 2003-371981/35.
 DR N-PSDB; ADC32599.
 DR
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Example 2; SEQ ID NO 3448; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC39919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting of
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human contig-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 145 AA;
 Query Match 42.9%; Score 63; DB 7; Length 145;
 Best Local Similarity 54.2%; Pred. No. 8.7;
 Matches 13; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 3 QGPGAGSGWEEGSGSPGVTPLFSP 26
 Db 71 QGDRGTGEGSGSGSPGPTAGMWP 94
 RESULT 12
 ABB11475
 ID ABB11475 standard; peptide; 575 AA.
 XX
 AC ABB11475;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human R31449_3 homologue, SEQ ID NO:1845.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW hematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW

KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 OS Homo sapiens.
 XX WO200157188-A2.
 PN XX
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US003800.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang Yt, Liu C, Drmanac RT;
 XX WPI; 2001-457740/49.
 DR N-PSDB; ABA08719.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 XX
 PS Claim 20; Page 197; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention
 XX Sequence 575 AA;

Query Match 41.8%; Score 61.5; DB 4; Length 575;
 Best Local Similarity 56.5%; Pred. No. 53;
 Matches 13; Conservative 2; Mismatches 7; Indels 1; Gaps 1;
 QY 1 GMOG-PAGSGWEGSGSPPGVTP 22
 DB 421 GGOGLPEGWGLEKGEGLPPGIPP 443
 RESULT 13
 AAB51865
 ID AAB51865 standard; protein; 98 AA.
 XX AC AAB51865;
 XX DT 16-FEB-2001 (first entry)
 XX DE Human secreted protein sequence encoded by gene 39 SEQ ID NO:98.
 XX KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 KW neutropic; neuroprotective; antibacterial; virucide; fungicide;
 KW ophthalmological; vulnery; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorders; cancer; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; nervous system disorder;
 KW Alzheimer's disease; ocular disorder; wound healing; skin aging.
 XX OS Homo sapiens.
 XX PN WO200061626-A1.
 XX PD 19-OCT-2000.
 XX PF 06-APR-2000; 2000WO-US009066.
 XX PR 09-APR-1999; 99US-0128698P.
 XX PR 20-JAN-2000; 2000US-0176926P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-619227/59.
 DR N-PSDB; AAC93517.
 XX
 PT New nucleic acid molecules encoding 49 human secreted proteins for
 PT diagnosing, preventing or ameliorating medical conditions and used for
 PT food additives or preservatives.
 XX
 PS Claim 11; Page 478; 516pp; English.
 CC Polynucleotide sequences AAC93479 - AAC93527 represent cDNA encoding
 CC human secreted proteins AAB51827 - AAB51875. Sequences AAB51876 -
 CC AAB51927 represent alternative polypeptides encoded by the genes, and
 CC amino acid sequences with which they share homology. The genes and
 CC proteins have activities dependent on the tissues and cells in which they
 CC are expressed. Examples of their activities include immunosuppressive;
 CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
 CC vasotropic; cerebroprotective; neutropic; neuroprotective; antibacterial;
 CC virucide; fungicide; ophthalmological; and vulnery. The secreted
 CC proteins, polynucleotides, antagonists and agonists may be useful in
 CC treating, preventing and/or diagnosing diseases and disorders such as
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC e.g. corneal infection. The polypeptides can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to

CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional
 CC components. Oligonucleotides AAC93470 - AAC93478 and peptide AAB51826 are
 CC used in the isolation and characterisation of the proteins and
 CC polynucleotides of the invention

XX SQ Sequence 98 AA;

Query Match 41.5%; Score 61; DB 3; Length 98;
 Best Local Similarity 59.1%; Pred. No. 10;
 Matches 13; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 5 PAGSGWEGSGSPGVTPLFSP 26
 DB 64 PRGSGWERAPGCP--VTPLTLP 83

RESULT 14
 AAR37744
 ID AAR37744 standard; protein; 54 AA.

AC AAR37744;

XX 25-MAR-2003 (revised)

DT 07-SEP-1993 (first entry)

XX Collagen-like polymer DCP- (DB)3.

XX Recombinant; collagen-like polymer; CLP; tripeptide; helix; membrane;
 KW fibre; film; coating; triad sequence; collagen; mammalian; moulding;
 KW hydrogel; interchain linkage; colloid suspension; DCP; antibody.

XX Synthetic.

XX WO9310154-A1.

PN 27-MAY-1993.

XX 04-NOV-1992; 92WO-US009485.

PR 12-NOV-1991; 91US-00791960.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

PI Cappello J, Ferrari FA;

XX WPI; 1993-182496/22.

DR N-PSDB; AAQ43035.

XX High mol. wt. collagen-like protein polymers - capable of being produced
 PT in unicellular microorganisms.

PS Disclosure; Page 52; 82pp; English.

CC The sequences given in AAR37744 and AAR37747 represent examples of
 CC recombinantly produced DCP collagen-like polymers (CLPs) which consist of
 CC repeated tripeptide sequences selected from a wide range of GXY
 CC sequences, where X and Y can be any amino acid. The DNA encoding these
 CC sequences can be cloned into plasmids and used to transform E. coli to
 CC produce the DCP proteins. DCP peptides comprise repeated units of: B =
 CC GSRGDPGPP, C = GAGGPGAPK and/or D = GAGGPGAGG. These polymers have
 CC molecular weights of >30 kD and are able to form helices due to
 CC tripeptide linkages. These polymers pref. contain a proportion of
 CC tripeptide triad sequences found in natural collagens, pref. mammalian
 CC collagens. The CLPs impart unique characteristics to materials such as
 CC fibres, membranes, films, coatings, hydrogels, colloid suspensions and
 CC moulded articles. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 54 AA;

Query Match 40.8%; Score 60; DB 2; Length 54;
 Best Local Similarity 57.9%; Pred. No. 7.3;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GMQGPAGSGWEGSGSPPG 19
 DB 1 GAQGPAGPGGSRGDPGPPG 19

RESULT 15

AAR93255
 ID AAR93255 standard; protein; 54 AA.

AC AAR93255;

XX 25-MAR-2003 (revised)

DT 24-FEB-1997 (first entry)

XX Collagen-like polymer sequence D4/D5 unit (DB)3.

XX collagen; repetitive triad motif; recombinant production; photographic;
 KW medical; structural; fibre.

XX Synthetic.

XX US5496712-A.

XX 05-MAR-1996.

XX 05-NOV-1992; 92US-00972032.

XX 06-NOV-1990; 90US-00609716.

PR 12-NOV-1991; 91US-00791960.

XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

PI Cappello J, Ferrari FA;

XX WPI; 1996-150728/15.

DR N-PSDB; AAT16769.

XX Collagen-like polymers comprising repetitive triads - produced in
 PT unicellular organisms with improved characteristics, useful in, e.g.
 PT photographic and medical fibres.

PS Example 3; Col 23-24; 43pp; English.

CC The invention concerns collagen-like polymers having repetitive triads
 CC with reduced proline content, and where glycine is the initial amino acid
 CC and the subsequent amino acids are varied. The choice of triads utilised
 CC in a recombinant collagen-like polymer are chosen in order to affect
 CC properties such as helix stability, hydration, solubility, gel point,
 CC biodegradation and immunogenicity. Triads of particular interest include
 CC GDR, GPA, GPP, GAS, GPG, GFS, GAG, GSP, GLO, GPR, GPK, GAK, GAR, GER,
 CC GDR, GEP, GDA, GAH and GEA. The collagen-like polymers may impart new
 CC characteristics, finding wide use in photographic, medical, structural
 CC and fibre applications, and are capable of being produced in unicellular
 CC microorganisms at high mol. wts. and in high efficiency. The present
 CC sequence, encoded by clone pMT0224 (see AAT16769), was identified to
 CC contain the sequenced gene 4 or 5 monomer sequence (DB)3. The Sequenced
 CC collagen-like polymers are used as immunogens for the prepn. of
 CC antibodies. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-
 CC MAR-2003 to correct PA field.)

XX SQ Sequence 54 AA;

Query Match 40.8%; Score 60; DB 2; Length 54;
 Best Local Similarity 57.9%; Pred. No. 7.3;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GMQGPAGSGWEGSGSPPG 19
 DB 1 GAQGPAGPGGSRGDPGPPG 19

Search completed: October 29, 2005, 01:03:13

Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 00:54:48 ; Search time 16 Seconds
(without alignments)
156.352 Million cell updates/sec

Title: US-09-980-263-1

Perfect score: 147
Sequence: 1 GMOGPAGSGWEGSGSPGVTLFSP 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	40.8	334	2 JC5980	ficolin-A precursor
2	60	40.8	1049	1 CGB07S	collagen alpha 1(I)
3	58.5	39.8	72	2 I33411	acetylcholinesterase
4	58.5	39.8	1076	2 J22217	major surface glycoprotein
5	58	39.5	591	2 A84765	hypothetical protein
6	58	39.5	888	2 S28791	collagen alpha 1(X)
7	58	39.5	1464	2 S59856	collagen alpha 1(I)
8	57	38.8	317	2 T19143	hypothetical protein
9	57	38.8	635	2 A57131	collagen alpha 2(I)
10	57	38.8	1366	1 CGH02S	collagen alpha 2(I)
11	56	38.1	282	1 JS0168	collagen col-8 - C
12	56	38.1	282	2 T16036	cuticular collagen
13	56	38.1	325	2 T18594	hypothetical protein
14	56	38.1	920	2 B34493	collagen alpha 1(I)
15	56	38.1	1373	1 A43291	collagen alpha 2(I)
16	56	38.1	3198	2 A3426	collagen alpha 2(I)
17	55.5	37.8	1712	1 CGH02B	collagen alpha 2(I)
18	55	37.4	316	2 T19288	hypothetical protein
19	55	37.4	439	2 D70954	hypothetical glycoprotein
20	55	37.4	469	2 A24450	collagen alpha 1(X)
21	55	37.4	633	2 B40983	collagen alpha 1(I)
22	55	37.4	886	2 IS0694	collagen alpha 1(I)
23	55	37.4	1453	2 S21626	collagen alpha 1(I)
24	55	37.4	1466	1 CGH07L	collagen type IV alpha 1(I)
25	55	37.4	1761	2 T13990	collagen type IV alpha 1(I)
26	54	36.7	316	2 T19291	hypothetical protein
27	54	36.7	1051	2 A35763	collagen alpha 2(I)
28	54	36.7	1142	2 JX0369	collagen alpha 1(X)
29	54	36.7	1418	2 T45467	collagen alpha 1(I)

30	54	36.7	1419	2 A41182	collagen alpha 1(I)
31	54	36.7	1487	1 CGH06C	collagen alpha 1(I)
32	54	36.7	1487	2 B41182	collagen alpha 1(I)
33	54	36.7	1549	2 T48103	type VII collagen
34	54	36.7	1573	2 S01845	DNA (cytosine-5')-DNA (cytosine-5')-DNA (cytosine-5')-DNA (cytosine-5')
35	54	36.7	1669	1 CGH04B	collagen alpha 1(I)
36	54	36.7	1669	1 CGH04B	collagen alpha 1(I)
37	53.5	36.4	608	2 A35021	myosin-light-chain
38	53	36.1	107	2 B61396	collagen alpha 1(I)
39	53	36.1	187	2 A35980	collagen alpha 1(I)
40	53	36.1	325	2 S02170	collagen alpha 1(I)
41	53	36.1	325	2 T16324	hypothetical protein
42	53	36.1	564	2 T49322	related to RNA-binding protein
43	53	36.1	921	2 S40495	collagen alpha 1(I)
44	53	36.1	921	2 S42617	collagen alpha 1(I)
45	53	36.1	931	2 S13580	collagen alpha 1(I)

ALIGNMENTS

RESULT 1

JC5980
ficolin-A precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: JC5980
R:Fujimori, Y.; Harumiya, S.; Fukumoto, Y.; Miura, Y.; Yagasaki, K.; Tachikawa, H.; Fujii, Biochem. Biophys. Res. Commun. 244, 796-800, 1998
A:Title: Molecular cloning and characterization of mouse ficolin-A.
A:Reference number: JC5980; MUID:98205801; PMID:9535745
A:Accession: JC5980
A:Molecule type: mRNA
A:Residues: 1-334 <FUJ>
A:Cross-references: UNIPROT:O70165; DDBJ:AB007813; NID:g2957011; PIDN:BAA25126.1; PID:dl
A:Experimental source: liver
C:Comment: This protein consists of both collagen- and fibrinogen- like domains.
F:1-21/Domain: signal sequence #status predicted <SIG>
F:50-64,68-106/Domain: collagen-like #status predicted <COL>
F:123-334/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 40.8%; Score 60; DB 2; Length 334;
Best Local Similarity 60.0%; Pred. No. 4.6;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 1 GMOGPAGSGWEGSGSPGV 20

Db 77 GMOGPAGSGWEGSGSPGV 96

RESULT 2

CGB07S
collagen alpha 1(III) chain - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C:Accession: A02862; A38001; A38002; A38003; A38004; A38005; S71946
R:Fietzek, P.P.; Allmann, H.; Kauterberg, J.; Henkel, W.; Wachter, E.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 809-820, 1979
A:Title: The covalent structure of calf skin type III collagen. I. The amino acid sequence of the triple helix.
A:Reference number: A02862; MUID:80026026; PMID:488906
A:Accession: A02862
A:Molecule type: protein
A:Residues: 1-242 <FIE>
A:Cross-references: UNIPROT:P04258
R:Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 821-832, 1979
A:Title: The covalent structure of calf skin type III collagen. II. The amino acid sequence of the triple helix.
A:Reference number: A38001; MUID:80026027; PMID:488907
A:Accession: A38001
A:Molecule type: protein
A:Residues: 243-422 <DEW1>
R:Bentz, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 833-840, 1979

A;Title: The covalent structure of calf skin type III collagen. III. The amino acid sequence
 A;Reference number: A38002; MUID:80026028; PMID:488908
 A;Accession: A38002
 A;Molecule type: protein
 A;Residues: 423-571 <BEN>
 R;Lang, H.; Glanville, R.W.; Fietzek, P.P.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 841-850, 1979
 A;Title: The covalent structure of calf skin type III collagen. IV. The amino acid sequence
 A;Reference number: A38003; MUID:80026029; PMID:488909
 A;Accession: A38003
 A;Molecule type: protein
 A;Residues: 572-808 <LAN>
 R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979
 A;Title: The covalent structure of calf skin type III collagen. V. The amino acid sequence
 A;Reference number: A38004; MUID:80026030; PMID:488910
 A;Accession: A38004
 A;Molecule type: protein
 A;Residues: 809-947 <DEW>
 R;Allmann, H.; Fietzek, P.P.; Glanville, R.W.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979
 A;Title: The covalent structure of calf skin type III collagen. VI. The amino acid sequence
 A;Reference number: A38005; MUID:80026031; PMID:488911
 A;Accession: A38005
 A;Molecule type: protein
 A;Residues: 948-1049 <ALL>
 R;Henkel, W.
 Biochem. J. 318, 497-503, 1996
 A;Title: Cross-link analysis of the C-telopeptide domain from type III collagen.
 A;Reference number: S71946; MUID:96404897; PMID:8809038
 A;Accession: S71946
 A;Molecule type: protein
 A;Residues: 87-106;1017-1029;1037-1049 <HEN>
 C;Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are highly conserved.
 C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline
 F;1-1049/Product: collagen alpha 1(I) chain #status experimental <CAB>
 F;1-14/Region: amino-terminal nonhelical telopeptide
 F;15-1040/Region: helical
 F;587-589/Region: cell attachment (R-G-D) motif
 F;752-754/Region: cell attachment (R-G-D) motif
 F;875-877/Region: cell attachment (R-G-D) motif
 F;878-880/Region: cell attachment (R-G-D) motif
 F;935-937/Region: cell attachment (R-G-D) motif
 F;1041-1049/Region: carboxyl-terminal nonhelical telopeptide
 F;95-107,119,938,950/Modified site: 5-hydroxylysine (Lys) #status experimental
 F;107,950/Modified site: allysine (Lys) #status predicted
 F;107/Binding site: carbohydrate (Lys) (covalent) #status experimental
 F;1040,1041/Disulfide bonds: interchain #status predicted

Query Match 40.8%; Score 60; DB 1; Length 1049;
 Best Local Similarity 57.1%; Pred. No. 15;
 Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GMOGPAGSGWEGSGSPPGVT 21
 Db 447 GPQGPAGNGETGPQGPQPT 467

RESULT 3
 153411
 acetylcholinesterase (EC 3.1.1.7), minor splice form - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
 C;Accession: I53411; R41573
 R;Karpel, R.; Ben Aziz-Alouy, R.; Sternfeld, M.; Ehrlich, G.; Ginzberg, D.; Tarroni, P.;
 Exp. Cell Res. 210, 268-277, 1994
 A;Title: Expression of three alternative acetylcholinesterase messenger RNAs in human tu
 A;Reference number: I53411; MUID:94131004; PMID:8299725
 A;Accession: I53411
 A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA
 A;Residues: 1-72 <KAR>
 A;Cross-references: GB:S71129; NID:9545224; PIDN:AAC60618.1; PID:9545225
 R;Li, Y.; Camp, S.; Rachinsky, T.L.; Getman, D.; Taylor, P.
 J. Biol. Chem. 266, 23083-23090, 1991
 A;Title: Gene structure of mammalian acetylcholinesterase. Alternative exons dictate tis
 A;Reference number: A41573; MUID:92078174; PMID:1744105
 A;Accession: A41573
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 30-46, 'R', 48-72 <LIY>
 A;Cross-references: GB:M76539; NID:9177976; PID:9553165
 C;Genetics:
 A;Gene: GDB:ACHE
 A;Cross-references: GDB:118746; OMIM:100740
 A;Map position: 7q22-7q22
 C;Keywords: alternative splicing; carboxylic ester hydrolase

Query Match 39.8%; Score 58.5; DB 2; Length 72;
 Best Local Similarity 50.0%; Pred. No. 1.4;
 Matches 14; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

Qy 1 GMOGPAGSGWEGSGS---PPGVTPPLFS 25
 Db 4 GMOGPAGSAGRGVGRQCNPSSLPLAS 31

RESULT 4
 JC2217
 major surface glycoprotein 5 - Pneumocystis carinii
 C;Species: Pneumocystis carinii
 C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
 C;Accession: JC2217
 R;Kitada, K.; Wada, M.; Nakamura, Y.
 DNA Res. 1, 57-66, 1994
 A;Title: Multi-gene family of major surface glycoproteins of Pneumocystis carinii: full-
 A;Reference number: JC2217; MUID:96051981; PMID:7584029
 A;Accession: JC2217
 A;Molecule type: mRNA
 A;Residues: 1-1076 <KIT>
 A;Cross-references: UNIPROT:Q01830; DBJ:D21827; NID:9425784; PIDN:BA04851.1; PID:dl005
 C;Superfamily: Pneumocystis carinii major surface glycoprotein MSG100
 C;Keywords: Glycoprotein
 F;245,471,574,804,837/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.8%; Score 58.5; DB 2; Length 1076;
 Best Local Similarity 65.0%; Pred. No. 23;
 Matches 13; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 3 GQGPAGSGWEGSGSPPGVTP 22
 Db 816 QAPAGS---SGSGSPPAVPP 832

RESULT 5
 A84765
 hypothetical protein At2g35130 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C;Accession: A84765
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: A84765
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-591 <STO>
 A;Cross-references: UNIPROT:O82178; GB:AE002093; NID:93668091; PIDN:AAC61823.1; GSPDB:GN
 C;Genetics:
 A;Gene: At2g35130

```

Query Match      39.5%; Score 58; DB 2; Length 888;
Best Local Similarity 53.8%; Pred. No. 22;
Matches 14; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

QY      1  GMQGPAGSGWEEGSGSPPGVTPLFSP 26
          |||||
DB       118 GAQGPAGLKGEGPGQPGPG--PVGSP 141
          |||||

```

A;Residues: 1-1464 <TOM>
A;Cross-references: UNIPROT:P08121; EMBL:X52046
R;Toman, D.
submitted to the EMBL Data Library, November 1994
A;Reference number: S62120
A;Accession: S62120
A;Molecule type: DNA
A;Residues: 1-866,'G',868-1464 <TOA>
A;Cross-references: EMBL:X52046; NID:G575321; PIDN:CAA36279.1; PID:G575322
R;Metaeranta, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.
Biochim. Biophys. Acta 1089, 241-243, 1991
A;Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
A;Reference number: S16176; MUID:91274355; PMID:2054384
A;Accession: S16373
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1442-1464 <MET>
A;Cross-references: EMBL:X57983; NID:G50476; PIDN:CAA41048.1; PID:G50477
C;Genetics:
A;Introns: 29/1; 95/3; 112/3; 150/3; 175/3; 193/3; 211/3; 229/3; 247/3; 265/3; 283/3; 2
58/3; 673/3; 706/3; 742/3; 760/3; 778/3; 796/3; 814/3; 850/3; 868/3; 886/3; 940/3; 976/
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix

```

F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-154/Domain: propeptide #status predicted <PRO>
F:32-92/Domain: von Willbrand factor type C repeat homology <VWC>
F:155-1464/Product: collagen alpha 1(III) chain #status predicted <MAT>
F:1236-1464/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match          39.5%;   Score 58;   DB 2;   Length 1464;
Best Local Similarity 52.4%;   Pred. NO. 37;
Matches 11;   Conservative 1;   Mismatches 9;   Indels 0;   Gaps 0;

QY      1  GMQGPAGSGWEEGSGSPPGVT 21
      1: ||||| | | | | | | |
Db      602  GLPGPAGKNGETPGQPGPPT 622

RESULT 8
T19143
hypothetical protein C0905.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19143
R:Palmer, S.
submitted to the EMBL Data Library, November 1994
A:Reference number: Z19080
A:Accession: T19143
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-317 <WIL>

```

Query Match	38.8%	Score 57	DB 2	Length 317
Best Local Similarity	57.9%	Pred. No. 10		
Matches	11	Conservative	2	Mismatches 6
				Indels 0
				Gaps 0

```

QY      1  GMQGPAGSGWEEGSGSPFG 19
      ||||| : : |||||
Db      211  GPQGPFGAGPQFGSGTFG 229
      ||||| : : |||||

RESULT 9
A57131 collagen alpha 2 (VII) chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: A57131
R:Muragaki, Y.; Jacenko, O.; Apté, S.; Mattei, M.G.; Ninomiya, Y.; Olsen, B.R.
J. Biol. Chem. 266, 7721-7727, 1991
A:Title: The alpha2(VII) collagen gene. A novel member of the short chain collagen family
A:Reference number: A57131; PMID:91210292; PMID:2019595
A:Accession: A57131
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-635 <MUR>

```

Query Match	38.8%;	Score 57;	DB 2;	Length 635;
Best Local Similarity	55.0%;	Pred. No. 21;		
Matches 11;	Conservative	1;	Mismatches .8;	Indels 0;
				Gaps 0;

A:Reference number: I54365; MUID:95187161; PMID:7881420
A:Accession: I68663
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 663-675, 'V', 677, 'P', 679-742, 'A', 744-746 <FOR>
A:Cross-references: GB:L47668; NID:g1009095; PIDN:AA59577.1; PID:g1009096
R:Niyibizi, C.; Bonadio, J.; Byers, P.H.; Eyre, D.R.
J. Biol. Chem. 267, 23108-23112, 1992
A:Title: Incorporation of type I collagen molecules that contain a mutant alpha 2(I) chain
A:Reference number: I55369; MUID:93054637; PMID:1385413
A:Accession: I55369
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 665-666, 'D', 668-670 <NIY>
A:Cross-references: GB:L00613; NID:g180888; PIDN:AA59384.1; PID:g180889
A:Note: mutant sequence from a patient with osteogenesis imperfecta
R:Bateman, J.F.; Hannigan, M.; Chan, D.; Cole, W.G.
Biochem. J. 276, 765-770, 1991
A:Title: Characterization of a type I collagen alpha 2(I) glycine-586 to valine substitution
A:Reference number: A56799; MUID:91291136; PMID:2064612
A:Accession: A56799
A:Molecule type: mRNA
A:Residues: 672-675, 'V', 677, 'P', 679-681 <BAT>
A:Cross-references: GB:S39878; NID:g1679911; PIDN:AA519314.1; PID:g232761
A:Note: sequence extracted from NCBI backbone (NCBIN:39878, NCBIP:39886)
A:Note: mutant sequence of patient with osteogenesis imperfecta type IV; the authors suggest
a method.
R:Maekelae, J.K.; Vuorio, T.; Vuorio, E.
Biochim. Biophys. Acta 1049, 171-176, 1990
A:Title: Growth-dependent modulation of type I collagen production and mRNA levels in culture
A:Reference number: S10768; MUID:90304220; PMID:2364107
A:Accession: S10768
A:Molecule type: mRNA
A:Residues: 960-1021, 'L', 1023-1188, 'D', 1190-1197, 'S', 1199-1356 <MAE>
A:Cross-references: EMBL:X55255; NID:g30101; PIDN:CAA39142.1; PID:g30102
A:Experimental source: fibroblast cell culture
R:Myers, J.C.; Chu, M.L.; Faro, S.H.; Clark, W.J.; Prockop, D.J.; Ramirez, F.
Proc. Natl. Acad. Sci. U.S.A. 78, 3516-3520, 1981
A:Title: Cloning a cDNA for the pro-alpha2 chain of human type I collagen.
A:Reference number: A18855; MUID:81273090; PMID:6267597
A:Accession: A18855
A:Molecule type: mRNA
A:Residues: 964-979, 'V', 981-1018, 'Q', 1020 <MYE>
A:Cross-references: GB:J00114; NID:g180393; PIDN:AAA51996.1; PID:g180394
A:Note: 1019-Leu was also found
R:Wenstrup, R.J.; Cohn, D.H.; Cohen, T.; Byers, P.H.
J. Biol. Chem. 263, 7734-7740, 1988
A:Title: Arginine for glycine substitution in the triple-helical domain of the products
A:Reference number: I55285; MUID:88227975; PMID:2897363
A:Accession: I55285
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1090-1107 <WEN1>
A:Cross-references: GB:M22816; NID:g179602; PIDN:AAA51844.1; PID:g179603
A:Accession: I70059
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1090-1101, 'R', 1103-1107 <WEN2>
A:Cross-references: GB:M22817; NID:g179606; PIDN:AAA51846.1; PID:g179607
A:Note: mutant sequence from a patient with osteogenesis imperfecta type IV
R:Myers, J.C.; Dickson, L.A.; de Wet, W.J.; Bernard, M.P.; Chu, M.L.; di Liberto, M.; Pe
J. Biol. Chem. 258, 10128-10135, 1983
A:Title: Analysis of the 3' end of the human pro-alpha-2(I) collagen gene. Utilization of
A:Reference number: S09175; MUID:83290853; PMID:6309769
A:Accession: S09175
A:Molecule type: DNA
Query Match 38.8%; Score 57; DB 1; Length 1366;
Best Local Similarity 47.6%; Pred. No. 46;
Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
OY 1 GMOGPAGSGWEEGSGSPPGVT 21

Db 766 GPNPGPGAGSGDGGPPGVT 786
RESULT 11
J50168
collagen col-8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: J50168
R:Cox, G.N.; Fields, C.; Kramer, J.M.; Rosenzweig, B.; Hirsch, D.
Gene 76, 331-344, 1989
A:Title: Sequence comparisons of developmentally regulated collagen genes of Caenorhabditis
A:Reference number: A91602; MUID:89326131; PMID:2753356
A:Accession: J50168
A:Molecule type: DNA
A:Residues: 1-282 <COX>
A:Cross-references: UNIPROT:P18833; GB:M25479; NID:g156267; PIDN:AAA27993.1; PID:g156268
A:Note: the authors translated the codon ACG for residue 43 as ARG
C:Comment: col-8, col-7 and col-19 belong to the same group of collagen genes.
C:Genetics:
A:Gene: col-8
A:Map position: III
A:Introns: 133/1; 259/3
F:95-124/Domain: helical <HX1>
F:141-269/Domain: helical <HX2>
Query Match 38.1%; Score 56; DB 2; Length 282;
Best Local Similarity 47.6%; Pred. No. 12;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
OY 1 GMOGPAGSGWEEGSGSPPGVT 21
Db 150 GLVGPGAGPGQGRHPPGPT 170
RESULT 12
T16036
cuticle collagen col-8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16036
R:Menezes, S.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid FilH8.
A:Reference number: Z18450
A:Accession: T16036
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-282 <MEN>
A:Cross-references: UNIPROT:P18833; EMBL:U40187; NID:g1055194; PID:g1055196; PIDN:AAA811
C:Genetics:
A:Gene: CESP:col-8
A:Introns: 133/1; 259/3
Query Match 38.1%; Score 56; DB 2; Length 282;
Best Local Similarity 47.6%; Pred. No. 12;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
OY 1 GMOGPAGSGWEEGSGSPPGVT 21
Db 150 GLVGPGAGPGQGRHPPGPT 170
RESULT 13
T18594
hypothetical protein AC3.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18594
R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z18995

A:Accession: T18594
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-325 <WIL>
 A:Cross-references: UNIPROT:O17402; EMBL:Z711177; PIDN:CAA94869.1; GSPDB:GN00023; CESP:AC
 A:Experimental source: clone AC3
 C:Genetics:
 A:Gene: CESP:AC3.6
 A:Map position: 5
 A:Introns: 47/3; 101/1; 270/3

Query Match 38.1%; Score 56; DB 2; Length 325;
 Best Local Similarity 57.9%; Pred. No. 14;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPGP 19
 DB 218 GSPGAGSGRGRGPGP 236

RESULT 14
 B34493
 collagen alpha 1(IX) chain long form precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 13-Jul-1990 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
 C:Accession: B34493; S21861; A28360; B28360; S22243; A20982; A28754; A30973; S22240
 R:Nishimura, I.; Murakami, Y.; Olsen, B.R.
 J. Biol. Chem. 264, 20033-20041, 1989
 A:Title: Tissue-specific forms of type IX collagen-proteoglycan arise from the use of two
 A:Reference number: A34493; MUID:90062114; PMID:2584206
 A:Accession: B34493
 A:Molecule type: mRNA
 A:Residues: 1-416 <NIS>
 A:Cross-references: UNIPROT:P12106; UNIPROT:Q90779; GB:M28659; NID:G211623; PIDN:AAA4870
 A:Note: the authors translated the codon TAC for residue 25 as Ile, GAA for residue 51
 &Olsen, B.R.
 submitted to the EMBL Data Library, December 1987

A:Reference number: S21861
 A:Accession: S21861
 A:Molecule type: mRNA
 A:Residues: 1-24, 'I', 26-50, 'Q', 52-299 <OLS>
 A:Cross-references: EMBL:J03539; NID:G211573; PIDN:AAA48702.1; PID:G211574
 R:Vasios, G.; Nishimura, I.; Konomi, H.; van der Rest, M.; Ninomiya, Y.; Olsen, B.R.
 J. Biol. Chem. 263, 2324-2329, 1988
 A:Title: Cartilage type IX collagen-proteoglycan contains a large amino-terminal globula
 A:Reference number: A28360; MUID:88115376; PMID:3339014
 A:Accession: A28360
 A:Molecule type: mRNA
 A:Residues: 1-24, 'I', 26-50, 'Q', 52-266 <VAS>
 A:Cross-references: EMBL:J03539
 A:Accession: B28360
 A:Molecule type: protein
 A:Residues: 31-38;49-50, 'Q', 52-54;97-114;152-158;259-266 <VA2>
 R:Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; McO
 maguchi, N.; Olsen, B.R.
 In Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pre
 A:Title: The molecular biology of collagens with short triple-helical domains.
 A:Reference number: S22243
 A:Accession: S22243
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-24, 'I', 26-50, 'Q', 52-920 <NI1>
 R:Ninomiya, Y.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 81, 3014-3018, 1984
 A:Title: Synthesis and characterization of cDNA encoding a cartilage-specific short coll
 A:Reference number: A20982; MUID:84221895; PMID:6328487
 A:Accession: A20982
 A:Molecule type: mRNA
 A:Residues: 183-920 <NI2>
 A:Cross-references: EMBL:K01702; NID:G211498; PIDN:AAA48675.1; PID:G211499
 R:Lozano, G.; Ninomiya, Y.; Thompson, H.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4050-4054, 1985
 A:Title: A distinct class of vertebrate collagen genes encodes chicken type IX collagen

A:Reference number: A94705; MUID:85216631; PMID:3858862
 A:Accession: A28754
 A:Molecule type: DNA
 A:Residues: 833-920 <LOZ>
 A:Cross-references: EMBL:M11324
 R:van der Rest, M.; Mayne, R.; Ninomiya, Y.; Seidah, N.G.; Chretien, M.; Olsen, B.R.
 J. Biol. Chem. 260, 220-225, 1985
 A:Title: The structure of type IX collagen.
 A:Reference number: A30973; MUID:85080081; PMID:2981204
 A:Accession: A30973
 A:Molecule type: protein
 A:Residues: 262-272;778-782, 'Q', 784-797 <VAN>
 R:Mayne, R.; van der Rest, M.; Ninomiya, Y.; Olsen, B.R.
 Ann. N. Y. Acad. Sci. 460, 38-46, 1985
 A:Title: The structure of type IX collagen.
 A:Reference number: S22238; MUID:86185164; PMID:3868958
 A:Accession: S22240
 A:Molecule type: protein
 A:Residues: 778-797 <MAY>
 C:Genetics:
 A:Introns: 5/2; 30/1; 56/1; 100/2; 232/3; 258/3; 265/3; 859/1
 C:Keywords: alternative splicing; cartilage; coiled coil; extracellular matrix; glycopro
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-920/Product: collagen alpha 1(IX) chain #status predicted <MAT>
 F:24-266/Domain: non-collagenous NC4 #status predicted <NC4>
 F:267-403/Domain: collagenous COL3 #status predicted <COL3>
 F:404-415/Domain: non-collagenous NC3 #status predicted <NC3>
 F:416-754/Domain: collagenous COL2 #status predicted <COL2>
 F:755-784/Domain: non-collagenous NC2 #status predicted <NC2>
 F:785-899/Domain: collagenous COL1 #status predicted <COL1>
 F:900-920/Domain: non-collagenous NC1 #status predicted <NC1>
 F:44-198,242-252/Disulfide bonds: #status predicted
 F:269,271,787,790,793,795/Modified site: hydroxyproline (Pro) #status experimental

Query Match 38.1%; Score 56; DB 2; Length 920;
 Best Local Similarity 42.3%; Pred. No. 40;
 Matches 11; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPGPVTPLFSP 26
 DB 617 GPRGPGSGRGPAGPGLFGKWP 642

RESULT 15
 A43291
 collagen alpha 2(I) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: A43291; A54328
 R:Phillips, C.L.; Morgan, A.L.; Lever, L.W.; Wenstrup, R.J.
 Genomics 13, 1345-1346, 1992
 A:Title: Sequence analysis of a full-length cDNA for the murine pro alpha 2(I) collagen
 A:Reference number: A43291; MUID:92372043; PMID:1505972
 A:Accession: A43291
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1373 <PHI>
 A:Cross-references: UNIPROT:Q01149; GB:X58251; NID:G50488; PIDN:CAA41205.1; PID:G50489
 A:Note: sequence extracted from NCBI backbone (NCBIP:112027)
 R:Phillips, C.L.; Lever, L.W.; Pinnell, S.R.; Quarles, L.D.; Wenstrup, R.J.
 J. Invest. Dermatol. 97, 980-984, 1991
 A:Title: Construction of a full-length murine Proalpha2(I) collagen cDNA by the polymera
 A:Reference number: A54328; MUID:92084969; PMID:1748823
 A:Accession: A54328
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-110 <PH2>
 C:Genetics:
 A:Gene: COL1A2
 C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
 F:1145-1373/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 38.1%; Score 56; DB 1; Length 1373;
 Best Local Similarity 48.0%; Pred. No. 61;
 Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 GMQGPAGSGWEEGSGSPPGVTPLFS 25
 Db 60 GPPGPGSGPGPGSPAPPGLTGNA 84

Search completed: October 29, 2005, 01:03:31
 Job time : 17 secs

RA
RT

```

NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6; TISSUE=Brain;
MEDLINE=23288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., DeRise J.G.,
Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Musny D.K., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettenman M., Madan A.A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M., Butterfield V.S.,
Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
SEQUENCE FROM N.A.
STRAIN=C57BL/6; TISSUE=Brain;
Strausberg R.;
Submitted (JAN-2004) to the ENBL/GenBank/DBBJ databases.
ENBL; BC065148; AAH5148.1; -.
GO; GO:0005737; Cytoplasm; IEA.
GO; GO:0006817; Phosphate transport; IEA.
InterPro; IPR001073; CIQ.
InterPro; IPR008160; Collagen.
InterPro; IPR008983; TNF_like.
Pfam; PF00386; CIQ; 1.
Pfam; PF01391; Collagen; 8.
PRINTS; PRO0007; COMPLEMENTC1Q.
SMART; SM00110; CIQ; 1.
PROSITE; PS01113; CIQ; 1.
Collagen.
SEQUENCE 699 AA; 66943 MW; FCEDA4FBBA4642646 CRC64;
Query Match 42.2%; Score 62; DB 2; Length 699;
Best Local Similarity 60.0%; Pred. No. 50;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps
QY 1 GMQGPGAGSGWEKGSPGPV 20
| : | | | | | | | | | |
Db 366 GRGPPGSGKEVGPGGPV 385
RESULT 7
ID Q910B9 PRELIMINARY; PRT; 1458 AA.
AC Q910B9;
DC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen a3(I).
GN NamesCOL1A3;
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
NCBI_TaxID=8022;
[1]
SEQUENCE FROM N.A.
MEDLINE=21257802; PubMed=113358497;
Saito M., Takenouchi Y., Kunisaki N., Kimura S.;
"Complete primary structure of rainbow trout type I collagen
consisting of al(I)a2(I)a3(I) heterotrimers.";
RT

```

```

RL Eur. J. Biochem. 268:2817-2827(2001).
DR EMBL; AB052836; BAB5662.1; -.
DR GO; GO:0005581; C:collagen; IEA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005201; F:extracellular matrix structural constituent; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR ProDom; PD000007; Clg_helix; 2.
DR SMART; SM002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VNC; 1.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS01184; VWFC_2; 1.
KW Collagen.
SQ SEQUENCE 1458 AA; 137757 MW; AB1F9F3410A98650 CRC64;

Query Match 42.2%; Score 62; DB 2; Length 1458;
Best Local Similarity 63.2%; Pred. No. 1.1e+02;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGV 19
Db 611 GPOGPAGSGQKGEGGPAG 629

RESULT 8
FCN1 MOUSE
ID FCN1 MOUSE STANDARD; PRT; 334 AA.
AC 070165;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ficolin 1 precursor (Collagen/fibrinogen domain-containing protein 1)
DE (Ficolin-A) (Ficolin A) (M-Ficolin).
GN Name=Fcni1; Synonyms=Fcna;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAALB/c; TISSUE=Liver;
RX MEDLINE=98205801; PubMed=9535745; DOI=10.1006/bbrc.1998.8344;
RA Fujimori Y., Harumiya S., Fukumoto Y., Miura Y., Yagasaki K.,
RA Tachikawa H., Fujimoto D.;
RT "Molecular cloning and characterization of mouse ficolin-A.";
RL Biochem. Biophys. Res. Commun. 244:796-800(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grummond J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Involved in serum exerting lectin activity. Binds GlcNAc
(CC (By similarity)).

```

```

CC -!- SUBUNIT: Homopolymer. Interacts with elastin (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Found on the monocyte surface (By
CC similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in liver and spleen.
CC -!- SIMILARITY: Belongs to the ficolin lectin family.
CC -!- SIMILARITY: Contains 1 collagenous domain.
CC -!- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB007813; BAA25126.1; -.
CC EMBL; BC019180; AAH19180.1; -.
CC PIR; JC5980; JC5980.
CC HSP; P02671; IFZD.
CC MGD; MGI:1340905; Fcna.
CC InterPro; IPR008161; Clg_helix.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR002181; Fibrinogen_C.
CC Pfam; PF01391; Collagen; 1.
CC Pfam; PF00147; Fibrinogen_C; 1.
CC ProDom; PD000007; Clg_helix; 1.
CC SMART; SM00186; FBG; 1.
CC PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
CC Collagen; Glycoprotein; Lectin; Multigene family; Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 334 Ficolin 1.
FT DOMAIN 50 88 Collagen-like.
FT DOMAIN 152 298 Fibrinogen C-terminal.
FT CARBOHYD 261 261 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 334 AA; 36298 MW; 9D30C05036AA04B1 CRC64;

Query Match 40.8%; Score 60; DB 1; Length 334;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGV 20
Db 77 GKWGPAGSGKGEGPTMGPPGV 96

RESULT 9
ID CAL3 BOVIN STANDARD; PRT; 1049 AA.
AC P04258;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Collagen alpha 1(III) chain.
GN Name=COL3A1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-242.
RX MEDLINE=80026025; PubMed=488906;
RA Fietzek P.P., Allmann H., Rautenberg J., Henkel W., Wachter E.,
RA Kuhn K.;
RT "The covalent structure of calf skin type III collagen. I. The amino
RT acid sequence of the amino terminal region of the alpha 1(III) chain
RT (positions 1-222).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:809-820(1979).
RN [2]
RP SEQUENCE OF 243-422.
RX MEDLINE=80026027; PubMed=488907;
RA Dewes H., Fietzek P.P., Kuhn K.;

```

Db 447 GFQGPAGKNGTGFQGPFGPT 467

RESULT 10

Q076271 PRELIMINARY; PRT; 904 AA.

AC Q076271;

DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Nongradient byssal.

DE Mytilus edulis (Blue mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;

OC Mytiloidea; Mytilidae; Mytilus.

OC NCBI_TaxID=6550;

[1]

RN SEQUENCE FROM N.A.

RP MEDLINE=98393676; PubMed=9724735; DOI=10.1073/pnas.95.18.10517;

RX Qin X.X., Waite J.H.;

RA "A potential mediator of collagenous block copolymer gradients in mussel byssal threads.;"

RL Proc. Natl. Acad. Sci. U.S.A. 95:10517-10522(1998).

DR EMBL; AF0433945; AAC33847.1; "-

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0006817; P:phosphate transport; IEA.

DR InterPro; IPR008161; Clg_helix.

DR InterPro; IPR008160; Collagen.

DR Pfam; PF01391; Collagen; 7.

DR ProDom; PD000007; Clg_helix; 1.

KW Collagen.

Q076271 SEQUENCE 904 AA; 77883 MW; 5529135651AD4C40 CRC64;

Query Match 40.5%; Score 59.5; DB 2; Length 904;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 13; Conservative 3; Mismatches 9; Indels 1;

Qy 1 GNGGPAGKNGTGFQGPFGPT 26

Db 227 GPRGPAGPGQGHGPGG-PPGHSP 251

RESULT 11

Q8MW55 PRELIMINARY; PRT; 905 AA.

AC Q8MW55;

DT 01-OCT-2002 (TReMBLrel. 22, Created)

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Precollagen-NG.

OC Mytilus galloprovincialis (Mediterranean mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;

OC Mytiloidea; Mytilidae; Mytilus.

OC NCBI_TaxID=29158;

[1]

RN SEQUENCE FROM N.A.

RP MEDLINE=22038007; PubMed=12042339;

RX Lucas J.M., Vaccaro E., Waite J.H.;

RA "A molecular, morphometric and mechanical comparison of the structural elements of byssus from Mytilus edulis and Mytilus galloprovincialis.;"

RL J. Exp. Biol. 205:1807-1817(2002).

DR EMBL; AF448524; AAM34599.1; "-

DR GO; GO:0005737; C:cytoplasm; IEA.

DR GO; GO:0006817; P:phosphate transport; IEA.

DR InterPro; IPR008160; Collagen.

DR Pfam; PF01391; Collagen; 7.

KW Collagen.

Q08MW55 SEQUENCE 905 AA; 79251 MW; 499F6F6E26A5C182 CRC64;

Query Match 40.5%; Score 59.5; DB 2; Length 905;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 13; Conservative 3; Mismatches 9; Indels 1;

DR

DR GO; GO:0005581; C:collagen; IEA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005201; F:extracellular matrix structural constituent; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR009041; PMP_SGCI.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 16.
DR Pfam; PF00093; VWC; 1.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF_C; 1.
DR PROSITE; PS0184; VWF_C; 2; 1.
KW Collagen.
SQ SEQUENCE 1447 AA; 137144 MW; 9CABD561F5BA36BF CRC64;
Query Match 40.1%; Score 59; DB 2; Length 1447;
Best Local Similarity 46.2%; Pred. No. 2.4e+02;
Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
QY 1 GMOGPAGSGWEGSGSPGCVTFLFSP 26
DB 607 GPAGPAGERGEQGAAGPPGFGGLPGP 632
RESULT 15
Q95LJ6 PRELIMINARY; PRT; 135 AA.
AC Q95LJ6;
*DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Osada N., Hida M., Kusuda J., Tanuma R., Hirata M., Suto Y., Hirai M.,
RA Terao K., Sugano S., Hashimoto K.;
RT "Cynomolgus monkey testicular cDNAs for discovery of novel human genes
in the human genome sequence."
RL BMC Genomics 3:36-36(2002).
DR EMBL; AB072792; BAB69761.1; -.
KW Hypothetical protein.
SQ SEQUENCE 135 AA; 14280 MW; 398CC2768FFA25D7 CRC64;
Query Match 39.8%; Score 58.5; DB 2; Length 135;
Best Local Similarity 45.5%; Pred. No. 24;
Matches 15; Conservative 1; Mismatches 8; Indels 9; Gaps 2;
QY 3 QGPAG-SGWEGSGSP-----PGVTPFLFSP 26
DB 82 RGPVAGSGWEGACGQPTSLTPFRVTPCQVP 114
Search completed: October 29, 2005, 01:04:35
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 01:01:09 ; Search time 41 Seconds
(without alignments)
47.338 Million cell updates/sec

Title: US-09-980-263-1
Perfect score: 147

Sequence: 1 GMOGPAGSGWEEGSGPPGVTLFSP 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp:*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp:*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp:*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp:*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS.COMB.pdp:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp:*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	100.0	26	2	US-08-370-156-27
2	147	100.0	31	2	US-08-370-156-11
3	147	100.0	31	3	US-08-990-085-23
4	147	100.0	31	3	US-08-975-084-3
5	147	100.0	31	4	US-08-380-532-13
6	147	100.0	600	2	US-08-370-156-4
7	147	100.0	600	3	US-08-814-095-4
8	147	100.0	600	3	US-08-975-084-1
9	140	95.2	25	3	US-08-975-084-2
10	60	40.8	54	1	US-07-972-032-59
11	60	40.8	54	1	US-08-642-255-71
12	60	40.8	93	1	US-07-972-032-60
13	60	40.8	93	1	US-07-972-032-61
14	60	40.8	633	1	US-08-642-255-73
15	60	40.8	1065	1	US-08-642-255-72
16	59	40.1	532	1	US-08-494-168-9
17	57.5	39.1	280	3	US-09-383-586-18
18	57.5	39.1	280	4	US-09-823-038A-18
19	57	38.8	520	4	US-09-252-991A-29451
20	57	38.8	1024	3	US-08-931-820-2
21	57	38.8	1366	3	US-08-963-825-19
22	57	38.8	1366	3	US-09-500-811-19
23	57	38.8	1366	3	US-09-570-573-19
24	57	38.8	1366	3	US-09-548-608-19
25	57	38.8	1366	4	US-09-585-887-10
26	57	38.8	1366	4	US-09-289-578-10
27	57	38.8	1366	4	US-09-949-016-5882

28 57 38.8 1442 2 US-08-316-650-12 Sequence 12, Appl
29 57 38.8 1442 5 PCT-US95-03251-12 Sequence 12, Appl
30 56 38.1 161 4 US-09-252-991A-17843 Sequence 17843, A
31 56 38.1 475 4 US-09-252-991A-26219 Sequence 26219, A
32 55.5 37.8 549 1 US-08-494-168-8 Sequence 8, Appl
33 55.5 37.8 1712 4 US-09-961-403-9 Sequence 9, Appl
34 55 37.4 101 3 US-08-990-571-68 Sequence 68, Appl
35 55 37.4 101 4 US-09-528-784A-68 Sequence 68, Appl
36 55 37.4 101 4 US-09-569-098A-68 Sequence 69, Appl
37 55 37.4 105 3 US-08-990-571-69 Sequence 69, Appl
38 55 37.4 105 4 US-08-528-784A-69 Sequence 69, Appl
39 55 37.4 105 4 US-09-569-098A-69 Sequence 69, Appl
40 55 37.4 504 4 US-09-252-991A-20694 Sequence 20694, A
41 55 37.4 1057 3 US-08-931-820-4 Sequence 4, Appl
42 55 37.4 1739 4 US-09-795-061-2 Sequence 2, Appl
43 54 36.7 18 1 US-07-972-032-84 Sequence 84, Appl
44 54 36.7 18 1 US-08-642-255-135 Sequence 135, App
45 54 36.7 30 3 US-08-845-258-47 Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-08-370-156-27
; Sequence 27, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mullford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-370-156-27

Query Match 100.0%; Score 147; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 7.7e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GMOGPAGSGWEEGSGPPGVTLFSP 26

Db 1 GMOGPAGSGWEEGSGPPGVTLFSP 26


```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,065
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/850,347
FILING DATE: 02-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,826
FILING DATE: 01-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00086
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-990-065-23

Query Match 100.0%; Score 147; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.2e-11;
Matches 26; Conservative 0; Mismatches 0; Indels

QY 1 GNOGPAGSGWEGSGPPGVTLFSP 26
DB 6 GNOGPAGSGWEGSGPPGVTLFSP 31
|||||
|||||

RESULT 4
US-08-975-084-3
Sequence 3, Application US/08975084
Patent No. 6258780
GENERAL INFORMATION:
APPLICANT: SOREQ, Hermona
APPLICANT: FRIEDMAN, Alon
APPLICANT: SEIDMAN, Shlomo
APPLICANT: KAUFER, Daniela
TITLE OF INVENTION: A METHOD AND COMPOSITION FOR ENABLING
TITLE OF INVENTION: PASSAGE THROUGH THE BLOOD-BRAIN-BARRIER
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Kohn & Associates
STREET: 30500 No. 6258780thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,084
FILING DATE: 11-NOV-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00082
TELECOMMUNICATION INFORMATION:

```

TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-975-084-3

Query Match 100.0%; Score 147; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.2e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGSPPGVTPLFSP 26
DB 6 GMOGPAGSGWEEGSGSPPGVTPLFSP 31

RESULT 5

US-09-380-532-13
Sequence 13, Application US/09380532
Patent No. 6475998

GENERAL INFORMATION:

APPLICANT: Soreq, Hermona
Seidman, Shlomo
Shohami, Esther

TITLE OF INVENTION: TREATMENT OF INJURY TO THE CENTRAL NERVOUS SYSTEM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kohn & Associates

STREET: 30500 No. 6475998thwestern Hwy.

CITY: Farmington Hills

STATE: Michigan

COUNTRY: US

ZIP: 48334

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,532

FILING DATE: 12-No. 6475998-1999

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Montgomery, Ilene N.

REGISTRATION NUMBER: 38,972

REFERENCE/DOCKET NUMBER: 2391.00089

TELECOMMUNICATION INFORMATION:

TELEPHONE: (248) 539-5050

TELEFAX: (248) 539-5055

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-380-532-13

Query Match 100.0%; Score 147; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.2e-11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGSPPGVTPLFSP 26
DB 6 GMOGPAGSGWEEGSGSPPGVTPLFSP 31

RESULT 6

US-08-370-156-4

Sequence 4, Application US/08370156

Patent No. 5932780

GENERAL INFORMATION:

APPLICANT: Soreq, Hermona

APPLICANT: Zakut, Haim

APPLICANT: Shani, Moshe

TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR

TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Reising, Ethington, Barnard & Perry

STREET: P.O. Box 4390

CITY: Troy

STATE: Michigan

COUNTRY: US

ZIP: 48099

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/370,156

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Kohn, Kenneth I.

REGISTRATION NUMBER: 30,955

REFERENCE/DOCKET NUMBER: P-307 (Malford)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810) 689-3500

TELEFAX: (810) 689-4071

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 600 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-370-156-4

Query Match 100.0%; Score 147; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGSPPGVTPLFSP 26
DB 575 GMOGPAGSGWEEGSGSPPGVTPLFSP 600

RESULT 7

US-08-814-095-4

Sequence 4, Application US/08814095

Patent No. 6025183

GENERAL INFORMATION:

APPLICANT: Soreq, Hermona

APPLICANT: Zakut, Haim

APPLICANT: Shani, Moshe

TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR

TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: KOHN & ASSOCIATES

STREET: 30500 No. 6025183thwestern Highway, Suite 410

CITY: Farmington Hills

STATE: Michigan

COUNTRY: U.S.

ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,095
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 600 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-814-095-4

Query Match 100.0%; Score 147; DB 3; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPLFSP 26
DB 575 GMOGPAGSGWEEGSGPPGVTPLFSP 600

RESULT 8
US-08-975-084-1
Sequence 1, Application US/08975084
Patent No. 6258780
GENERAL INFORMATION:
APPLICANT: SOREQ, Hermona
APPLICANT: FRIEDMAN, Alon
APPLICANT: SEIDMAN, Shlomo
APPLICANT: KAUFER, Daniela
TITLE OF INVENTION: A METHOD AND COMPOSITION FOR ENABLING
PASSAGE THROUGH THE BLOOD-BRAIN-BARRIER
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6258780thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,084
FILING DATE: 11-NOV-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00082
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 600 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-975-084-1

Query Match 100.0%; Score 147; DB 3; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPLFSP 26
DB 575 GMOGPAGSGWEEGSGPPGVTPLFSP 600

RESULT 9
US-08-975-084-2
Sequence 2, Application US/08975084
Patent No. 6258780
GENERAL INFORMATION:
APPLICANT: SOREQ, Hermona
APPLICANT: FRIEDMAN, Alon
APPLICANT: SEIDMAN, Shlomo
APPLICANT: KAUFER, Daniela
TITLE OF INVENTION: A METHOD AND COMPOSITION FOR ENABLING
PASSAGE THROUGH THE BLOOD-BRAIN-BARRIER
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6258780thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,084
FILING DATE: 11-NOV-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Montgomery, Ilene N.
REGISTRATION NUMBER: 38,972
REFERENCE/DOCKET NUMBER: 2391.00082
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-975-084-2

Query Match 95.2%; Score 140; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.8e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPLFSP 25
DB 1 GMOGPAGSGWEEGSGPPGVTPLFSP 25

RESULT 10
US-07-972-032-59

Sequence 59, Application US/07972032
Patent No. 5496712
GENERAL INFORMATION:
APPLICANT: Cappello, Joseph
APPLICANT: Ferrari, Franco A.
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE
TITLE OF INVENTION: PROTEIN POLYMERS
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bertram I. Rowland
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: CA 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/972,032
FILING DATE: 19921105
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/791,960
FILING DATE: 12-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-55556-1/BIR;PROP-08-1
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 59:
LENGTH: 54 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-972-032-59
Query Match 40.8%; Score 60; DB 1; Length 54;
Best Local Similarity 57.9%; Pred. No. 2.2;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 GMOGPAGSGWEGSGSPPG 19
DB 1 GAQGPAGGSGRGGDPGPG 19
RESULT 11
US-08-642-255-71
Sequence 71, Application US/08642255
Patent No. 5773249
GENERAL INFORMATION:
APPLICANT: CAPPELLO, Joseph
APPLICANT: FERRARI, Franco A.
TITLE OF INVENTION: High Molecular Weight Collagen-Like
TITLE OF INVENTION: Protein Polymers
NUMBER OF SEQUENCES: 135
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HORBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROWLAND, Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A55556-3/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-642-255-71
Query Match 40.8%; Score 60; DB 1; Length 54;
Best Local Similarity 57.9%; Pred. No. 2.2;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 GMOGPAGSGWEGSGSPPG 19
DB 1 GAQGPAGGSGRGGDPGPG 19
RESULT 12
US-07-972-032-60
Sequence 60, Application US/07972032
Patent No. 5496712
GENERAL INFORMATION:
APPLICANT: Cappello, Joseph
APPLICANT: Ferrari, Franco A.
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT COLLAGEN-LIKE
TITLE OF INVENTION: PROTEIN POLYMERS
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bertram I. Rowland
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: CA 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/972,032
FILING DATE: 19921105
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/791,960
FILING DATE: 12-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-55556-1/BIR;PROP-08-1
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

```

RESULT 14
US-08-642-255-73
; Sequence 73, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-255-73

Query Match 40.8%; Score 60; DB 1; Length 633;
Best Local Similarity 57.9%; Pred. NO. 26;
Matches 11; Conservative 0; Mismatches 8; Indels

QY 1 GMOGPAGSGWBERGSGSPPG 19
Db 52 GAQGPAGGSGRDPGPG 70

RESULT 15
US-08-642-255-72
; Sequence 72, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROWLAND, Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A55556-3/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1065 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-642-255-72

Query Match 40.8%; Score 60; DB 1; Length 1065;
Best Local Similarity 57.9%; Pred. No. 44;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GNOGPAGSGWEGSGSPPG 19
Db 52 GAQGPAGGGRGDPGPPG 70
```

Search completed: October 29, 2005, 01:05:21
Job time : 43 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2005, 01:03:40 ; Search time 165 Seconds

(without alignments)
65.873 Million cell updates/sec

Title: US-09-980-263-1

Perfect score: 147

Sequence: 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09A_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	26	10	US-09-998-042-1
2	147	100.0	37	10	US-09-998-042-6
3	147	100.0	53	10	US-09-998-042-7
4	68	46.3	105	16	US-10-425-115-356384
5	61.5	41.8	157	16	US-10-425-115-274774
6	61.5	41.8	575	15	US-10-276-774-1845
7	61	41.5	138	16	US-10-425-115-321226
8	60	40.8	334	18	US-10-820-155-133
9	60	40.8	334	18	US-10-820-155-133
10	60	40.8	528	16	US-10-425-115-250621
11	60	40.8	534	15	US-10-425-114-70635

12	60	40.8	1466	15	US-10-402-089-4	Sequence 4, Appli
13	60	40.8	1466	15	US-10-402-089-6	Sequence 6, Appli
14	60	40.8	1466	15	US-10-402-072A-4	Sequence 4, Appli
15	60	40.8	1466	15	US-10-402-072A-6	Sequence 6, Appli
16	59	40.1	78	15	US-10-424-599-194897	Sequence 194897,
17	59	40.1	102	16	US-10-425-115-250617	Sequence 250617,
18	59	40.1	155	16	US-10-767-701-60190	Sequence 60190, A
19	59	40.1	598	16	US-10-425-115-250618	Sequence 250618,
20	59	40.1	603	16	US-10-437-963-167449	Sequence 167449,
21	58.5	39.8	72	18	US-10-756-149-4761	Sequence 4761, Ap
22	58.5	39.8	136	16	US-10-425-115-271420	Sequence 271420,
23	58	39.5	185	16	US-10-437-963-121283	Sequence 121283,
24	58	39.5	185	16	US-10-437-963-182893	Sequence 182893,
25	58	39.5	226	15	US-10-104-047-2316	Sequence 2316, Ap
26	58	39.5	251	15	US-10-094-749-2497	Sequence 2497, Ap
27	58	39.5	393	14	US-10-210-951-34	Sequence 34, Appl
28	58	39.5	393	14	US-10-211-884-34	Sequence 34, Appl
29	58	39.5	393	15	US-10-211-858-34	Sequence 2518, Ap
30	58	39.5	393	15	US-10-094-749-2518	Sequence 2, Appli
31	58	39.5	393	20	US-11-052-503-2	Sequence 108, App
32	58	39.5	638	13	US-10-001-887-108	Sequence 42, Appl
33	58	39.5	638	15	US-10-138-588-42	Sequence 108, App
34	58	39.5	638	17	US-10-958-863-108	Sequence 4, Appli
35	58	39.5	703	14	US-10-219-449-4	Sequence 4, Appli
36	58	39.5	703	17	US-10-958-858-4	Sequence 2, Appli
37	58	39.5	717	14	US-10-219-449-2	Sequence 2, Appli
38	58	39.5	717	17	US-10-958-858-2	Sequence 40, Appl
39	58	39.5	733	15	US-10-138-588-40	Sequence 18, Appl
40	57.5	39.1	280	9	US-09-823-038A-18	Sequence 207493,
41	57	38.8	103	16	US-10-425-115-207493	Sequence 257391,
42	57	38.8	133	16	US-10-425-115-257391	Sequence 168962,
43	57	38.8	386	16	US-10-437-963-168962	Sequence 55894, A
44	57	38.8	407	15	US-10-425-114-55894	Sequence 979, App
45	57	38.8	674	9	US-09-925-299-379	

ALIGNMENTS

RESULT 1
US-09-998-042-1
; Sequence 1, Application US/09998042
; Publication No. US20030036632A1
; GENERAL INFORMATION:
; APPLICANT: YISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW
; TITLE OF INVENTION: ACETYLCHOLINESTERASE-DERIVED PEPTIDE AND USES THEREOF
; FILE REFERENCE: 7811/WO/99
; CURRENT APPLICATION NUMBER: US/09/998,042
; CURRENT FILING DATE: 2002-07-02
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PPT
; ORGANISM: HOMO SAPIENS
US-09-998-042-1
Query Match 100.0%; Score 147; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26
DB 1 GMOGPAGSGWEEGSGPPGVTPFLFSP 26

RESULT 2
US-09-998-042-6
; Sequence 6, Application US/09998042
; Publication No. US20030036632A1
; GENERAL INFORMATION:
; APPLICANT: YISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW
; TITLE OF INVENTION: ACETYLCHOLINESTERASE-DERIVED PEPTIDE AND USES THEREOF

```

; LOCATION: (1)::. (105)
; OTHER INFORMATION: unsure at all Xaa locations

```


OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-276-774-1845

Query Match 41.8%; Score 61.5; DB 15; Length 575;
Best Local Similarity 56.5%; Pred. No. 88;
Matches 13; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 GMOG-PAGSGWEGSGSPGPVTP 22
DB 421 GGQGLPEGWLEKGGELPPGIPP 443

RESULT 7

US-10-425-115-221226
; Sequence 221226, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 221226
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Zea mays
; NAME/KEY: unsure
; LOCATION: (1)-(138)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_133350C.1.pep
; US-10-425-115-221226

Query Match 41.5%; Score 61; DB 16; Length 138;
Best Local Similarity 52.4%; Pred. No. 26;
Matches 11; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 4 GPAGS--GWEEGSGSPGPVTP 22
DB 95 GPMGTGLGWRPCAGGPPGPIP 115

RESULT 8

US-10-820-155-131
; Sequence 131, Application US/10820155
; Publication No. US20050137126A1
; GENERAL INFORMATION:
; APPLICANT: NatImmune A/S
; APPLICANT: Wellguny, Dietmar
; APPLICANT: Jensenius, Jens Christian
; APPLICANT: Kongerslev, Leif
; APPLICANT: Mathiesen, Finn
; TITLE OF INVENTION: Treatment of SARS in individuals
; FILE REFERENCE: P 774 US00
; CURRENT APPLICATION NUMBER: US/10/820,155
; CURRENT FILING DATE: 2004-04-08
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-820-155-131

Query Match 40.8%; Score 60; DB 18; Length 334;
Best Local Similarity 60.0%; Pred. No. 78;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGV 20
DB 77 GKMGPAKGEPGTMGPPGV 96

RESULT 9

US-10-820-155-133
; Sequence 133, Application US/10820155
; Publication No. US20050137126A1
; GENERAL INFORMATION:
; APPLICANT: NatImmune A/S
; APPLICANT: Wellguny, Dietmar
; APPLICANT: Jensenius, Jens Christian
; APPLICANT: Kongerslev, Leif
; APPLICANT: Mathiesen, Finn
; TITLE OF INVENTION: Treatment of SARS in individuals
; FILE REFERENCE: P 774 US00
; CURRENT APPLICATION NUMBER: US/10/820,155
; CURRENT FILING DATE: 2004-04-08
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-820-155-133

Query Match 40.8%; Score 60; DB 18; Length 334;
Best Local Similarity 60.0%; Pred. No. 78;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEGSGSPPGV 20
DB 77 GKMGPAKGEPGTMGPPGV 96

RESULT 10

US-10-425-115-250621
; Sequence 250621, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 250621
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_160154C.1.pep
; US-10-425-115-250621

Query Match 40.8%; Score 60; DB 16; Length 528;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 14; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

QY 1 GMOGPA--GSGWEGSGSPGPVTPLFSP 26
DB 7 GRAGAIKRGWKYGSFVDGVFPVLPSP 34

RESULT 11

US-10-425-114-70635
; Sequence 70635, Application US/10425114
; Publication No. US20040034888A1

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70635
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73046D03_FLI.pep
US-10-425-114-70635

Query Match 40.8%; Score 60; DB 15; Length 534;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 14; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

Qy 1 GMOGPAGSGWEGSGSPPGVT 26
Db 13 GRAGGAIKRGWKYSGFVDGVFPVLP 40

*RESULT 12
US-10-402-089-4
; Sequence 4, Application US/10402089
; Publication No. US20040005663A1
; GENERAL INFORMATION:
; APPLICANT: Bell, Marcum P.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Polarek, James W.
; APPLICANT: Seeley, Todd W.
; TITLE OF INVENTION: PORCINE COLLAGENS AND GELATINS
; FILE REFERENCE: FP0402.3 CON
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US 09/709,700
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Bos Taurus
US-10-402-089-4

Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GMOGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPPT 623

RESULT 13
US-10-402-089-6
; Sequence 6, Application US/10402089
; Publication No. US20040005663A1
; GENERAL INFORMATION:
; APPLICANT: Bell, Marcum P.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Polarek, James W.
; APPLICANT: Seeley, Todd W.
; TITLE OF INVENTION: PORCINE COLLAGENS AND GELATINS
; FILE REFERENCE: FP0402.3 CON
```

```
; CURRENT APPLICATION NUMBER: US/10/402,089
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US 09/709,700
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-402-089-6

Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GMOGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPPT 623

RESULT 14
US-10-402-072A-4
; Sequence 4, Application US/10402072A
; Publication No. US20040018592A1
; GENERAL INFORMATION:
; APPLICANT: Bell, Marcum P.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Polarek, James W.
; APPLICANT: Seeley, Todd W.
; TITLE OF INVENTION: BOVINE COLLAGENS AND GELATINS
; FILE REFERENCE: FP0402.2 CON
; CURRENT APPLICATION NUMBER: US/10/402,072A
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US 09/709,700
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Bos Taurus
US-10-402-072A-4

Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 GMOGPAGSGWEGSGSPPGVT 21
Db 603 GPQGPAGKNGTGPQGPPT 623

RESULT 15
US-10-402-072A-6
; Sequence 6, Application US/10402072A
; Publication No. US20040018592A1
; GENERAL INFORMATION:
; APPLICANT: Bell, Marcum P.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Polarek, James W.
; APPLICANT: Seeley, Todd W.
; TITLE OF INVENTION: BOVINE COLLAGENS AND GELATINS
; FILE REFERENCE: FP0402.2 CON
; CURRENT APPLICATION NUMBER: US/10/402,072A
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US 09/709,700
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 1466
; TYPE: PRT
```

! ORGANISM: Sus scrofa
US-10-402-072A-6

Query Match 40.8%; Score 60; DB 15; Length 1466;
Best Local Similarity 57.1%; Pred.No. 3.1e+02;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GMOGPAGSGWEEGSGSPPGVT 21
| | | | | | | | | | | | | | | | | | | | | |
Db 603 GPQGPAGKNGETGPGPGPT 623

Search completed: October 29, 2005, 01:18:08
Job time : 166 secs